GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2003, 15:02:10 ; Search time 72 Seconds
                                                    (without alignments)
                                                    2713.136 Million cell updates/sec

Title: US-09-494-297-2
Perfect score: 3945
Sequence: 1 MKKTRFPNKLNTLNTQRVLS.............IAGISLGIWGIHTIRIRKHD 757

Scoring table: BLOSUM62
Scoring: Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SPTREMBL_23:*
             1: sp_archea:*
             2: sp_bacteria:*
             3: sp_fungi:*
             4: sp_human:*
             5: sp_invertebrate:*
             6: sp_mammal:*
             7: sp_mhc:*
             8: sp_organelle:*
             9: sp_phage:*
            10: sp_plant:*
            11: sp_rodent:*
            12: sp_virus:*
            13: sp_vertebrate:*
            14: sp_unclassified:*
            15: sp_rvirus:*
            16: sp_bacteriap:*
            17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3941	99.9	757	2	Q8GRA2	Q8gra2 streptococc
2	2058.5	52.2	756	2	Q8RP53	Q8rp53 streptococc
3	1965.5	49.8	742	2	Q92B47	Q92b47 streptococc
4	1862	47.2	744	16	Q8K8U7	Q8k8u7 streptococc
5	1254.5	31.8	524	16	Q8P2W3	Q8p2w3 streptococc
6	633.5	16.1	659	2	Q54953	Q54953 streptococc
7	633	16.0	698	2	Q8RP54	Q8rp54 streptococc
8	611	15.5	685	2	O33709	O33709 streptococc
9	603	15.3	638	2	O01924	O01924 streptococc
10	523	13.3	580	2	O47942	O47942 streptococc
11	272.5	6.9	106	2	O33715	O33715 streptococc
12	262.5	6.7	104	2	O33714	O33714 streptococc
13	241.5	6.1	108	2	O33711	O33711 streptococc
14	236.5	6.0	340	16	Q9A1S2	Q9a1s2 streptococc
15	224.5	5.7	696	16	Q8P2V7	Q8p2v7 streptococc
16	223	5.7	733	2	Q8G9G1	Q8g9g1 streptococc

17	222.5	5.6	116	2	O33721	O33721 streptococc
18	211.5	5.4	92	2	O33718	O33718 streptococc
19	211.5	5.4	103	2	O33712	O33712 streptococc
20	210.5	5.3	102	2	O33719	O33719 streptococc
21	209.5	5.3	102	2	O33716	O33716 streptococc
22	205.5	5.2	93	2	O33713	O33713 streptococc
23	204.5	5.2	102	2	O33720	O33720 streptococc
24	202.5	5.1	93	2	O33710	O33710 streptococc
25	202.5	5.1	1368	16	Q93M90	Q93m90 clostridium
26	185	4.7	84	2	O33717	O33717 streptococc
27	180	4.6	597	2	P72416	P72416 streptococc
28	178	4.5	898	2	O85081	O85081 moraxella c
29	176	4.5	905	2	O85088	O85088 moraxella c
30	174.5	4.4	298	2	Q93ED6	Q93ed6 streptococc
31	167	4.2	344	16	Q8K8U5	Q8k8u5 streptococc
32	164	4.2	1039	2	P72534	P72534 streptococc
33	161	4.1	1160	2	Q8RM86	Q8rm86 streptococc
34	161	4.1	1386	16	Q92DL0	Q92dl0 listeria in
35	159	4.0	1243	2	Q48588	Q48588 lactococcus
36	159	4.0	1612	16	Q8Y591	Q8y591 listeria mo
37	159	4.0	1681	16	Q8T1G8	Q8t1g8 dictyosteli
38	158.5	4.0	3269	2	Q9RGN5	Q9rgn5 lactobacill
39	158	4.0	656	2	Q8K8U1	Q8k8u1 streptococc
40	157.5	4.0	5005	16	Q9PP25	Q9pp25 ureaplasma
41	156	4.0	1530	16	Q8Y479	Q8y479 listeria mo
42	155	3.9	832	16	Q8YAF5	Q8yaf5 listeria mo
43	155	3.9	957	2	Q9AHL1	Q9ahl1 borrelia bu
44	153	3.9	993	5	Q8ILA1	Q8ila1 plasmodium
45	153	3.9	1011	2	Q9AHL0	Q9ahl0 borrelia bu

ALIGNMENTS

RESULT 1
Q8GRA2
ID   Q8GRA2
AC   Q8GRA2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Protein F-homologous protein.
GN   PFHP.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=291;
RA   Miyoshi-Akiyama T., Wakisaka N., Zhao J., Uchiyama T.;
RT   "Characterizaition of Protein F-homologous protein of M1-type group A
RT   streptococcus.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB083107; BAC20340.1; -.
SQ   SEQUENCE 757 AA; 363414E6FF078DFD CRC64;

Query Match   99.9%; Score 3941; DB 2; Length 757;
Best Local Similarity   99.9%; Pred. No. 1.1e-209;
Matches 756; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MKKTRFPNKLNTLNTQRVLSKNSKRFTVTLLVGVFLMIFALVTSMVGAKTVFGLIVESSTPN 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKKTRFPNKLNTLNTQRVLSKNSKRFTVTLLVGVFLMIFALVTSMVGAKTVFGLIVESSTPN 60

Qy   61 AINPDSSSEYRWYGYESYVRGHPYYKQFRVAHDLRVNLEGSSRSYQVYCFNLKKAFPLGSD 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AINPDSSSEYRWYGYESYVRGHPYYKQFRVAHDLRVNLEGSSRSYQVYCFNLKKAFPLGSD 120

Qy   121 SSVKKWYKKHDGISTKFEDYAMSPRITGDELNQKLRAVMYNGHPQNANGIMEGLEPLNAI 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 SSVKKWYKKHDGISTKFEDYAISPRITGDELNQKLRAVMYNGHPQNANGIMEGLEPLNAI 180

```
Qy    181  RVTQEAVWYYSDNAPISNPDESFKRESESNLVSTSQLSLMRQALKQLIIDPNLATKMPKQV  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  RVTQEAVWYYSDNAPISNPDESFKRESESNLVSTSQLSLMRQALKQLIIDPNLATKMPKQV  240

Qy    241  PDDFQLSIFESEDKGDRKYNKGYQNLLSGGLVPTKPPTPGDPPMPPNQPQTTSVLIRKYAI  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  PDDFQLSIFESEDKGDRKYNKGYQNLLSGGLVPTKPPTPGDPPMPPNQPQTTSVLIRKYAI  300

Qy    301  GDYSKLLEGATLQLTGDNVNSFQARVFSSNNDIGERIELSDGTYTLTELNSPAGYSIAEPI  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  GDYSKLLEGATLQLTGDNVNSFQARVFSSNNDIGERIELSDGTYTLTELNSPAGYSIAEPI  360

Qy    361  TFKVEAGKVYTIIDGKQIENPNKEIVEPYSVEAYNDFEEFSVLTTQNYAKFYYAKNKNGS  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TFKVEAGKVYTIIDGKQIENPNKEIVEPYSVEAYNDFEEFSVLTTQNYAKFYYAKNKNGS  420

Qy    421  SQVVYCFNADLKSPPDSEDGGKTMTPDFTTGEVKYTHIAGRDLFKYTVKPRDTDPDDFLK  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  SQVVYCFNADLKSPPDSEDGGKTMTPDFTTGEVKYTHIAGRDLFKYTVKPRDTDPDDFLK  480

Qy    481  HIKKVIEKGYREKGQAIEYSGLTETQLRAATQLAIYFTDSAELDKDKLKDYHGFGDMND  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  HIKKVIEKGYREKGQAIEYSGLTETQLRAATQLAIYFTDSAELDKDKLKDYHGFGDMND  540

Qy    541  STLAVAKILVEYAQDSNPPQLTDLDFFIPNNNKYQSLIGTQWHPEDLVDIIRMEDKKEVI  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  STLAVAKILVEYAQDSNPPQLTDLDFFIPNNNKYQSLIGTQWHPEDLVDIIRMEDKKEVI  600

Qy    601  PVTHNLTLRKTVTGLAGDRTKDFHFEIELKNNKQELLSQTVKTDKINLEFKDGKATINLK  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  PVTHNLTLRKTVTGLAGDRTKDFHFEIELKNNKQELLSQTVKTDKINLEFKDGKATINLK  660

Qy    661  HGESLTLQGLPEGYSYLVKETDSEGYKVKVNSQEVANATVSKTGITISDETLAFENNKEPV  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  HGESLTLQGLPEGYSYLVKETDSEGYKVKVNSQEVANATVSKTGITISDETLAFENNKEPV  720

Qy    721  VPTGVDQKINGYLALIVIAGISLGIWGIHTIRIKHD  757
           |||||||||||||||||||||||||||||||||||
Db    721  VPTGVDQKINGYLALIVIAGISLGIWGIHTIRIKHD  757
```

RESULT 2

```
Q8RP53
ID    Q8RP53            PRELIMINARY;      PRT;    756 AA.
AC    Q8RP53;
DT    01-JUN-2002 (TrEMBLrel. 21, Created)
DT    01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT    01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE    Cpa.
DE    CPA.12.
GN    Cpa.
OS    Streptococcus pyogenes.
OC    Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC    Streptococcus.
OX    NCBI_TaxID=1314;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=A735;
RX    MEDLINE=21843101; PubMed=11854196;
RA    Bessen D.E., Kalia A.;
RT    "Genomic localization of a T serotype locus to a recombinatorial
RT    zone encoding extracellular matrix-binding proteins in Streptococcus
RT    pyogenes.";
RL    Infect. Immun. 70:1159-1167(2002).
DR    EMBL; AF447492; AAL86406.1; -.
SQ    SEQUENCE    756 AA;  85698 MW;  88EDE087714EC4C464 CRC64;
```

```
Qy     73  YGVESYVRGHPIYXQFRVAHDLRVNLEGGSRSYQYYCFNLKKAFPLGSDSSVKKWYTKHDG  132
           |||:|| ||||||:|||| :|||||||||||||||||||| :| |||:|: || 
Db     73  YGYDSYDSSHPYYERFKVAHDLRVNLNGSSKSYQVYCFNINSHYPNRKNAFSKQWFKRVDG  131

Qy    133  ISTRFEDYAMSPRITGDELNQKLRAVMYNGHPQNANGIMEGLEPLNAIRVTQEAVWYYSD  192
           ||:||| |||||||:|||:| ||:||||||:||||||||||||:||||| || |||||||
Db    132  TGDVFTNYAQTPKIRGESLNNKLLSIMYNAYPKNANGYMDKIEPLNALLVTQQAVWYYSD  191

Qy    193  NAPISNPDESFKRESESNLVSTSQLSLMRQALKQLIIDPNLATKMPKQVPDDFQLSIFESE  252
           :: |||||||||||||||||||||||||||||||||||||||||| ::| :|||||:
Db    192  SS-YGNIKTLWASELKDGKIDFEQVKLMREAYSKLIISDDLEEFSKNKLPQGSKLNIFVPQ  250

Qy    253  DKGDRKYNKGYQNLLSGGLVPTKPPTPGDPPMPPNQPQTTSVLIRKYAIGDYSKLLEGATL  312
           |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db    251  D------KSVQNLLSAEYVPESPPAPGQSPEPVQTKKITSVIIRKYAEGDYSKLLEGATL  304

Qy    313  QLTGDNVNSFQARVFSSNNDIGERIELSDGTYTLTELNSPAGYSIAEPITFKVEAGKYYTI  372
           ||||||||||||| :| :||||||||||||||||||||||||| |||||||| | |||||
Db    305  RLTGEDILDFQEKVFQSNGTGEKIELSNGTYTLTETSSPDGYKIAEPIKFRVVNKKFVVI  364

Qy    373  -IDGKQIENPNKEIVEPYSVEAYNDFEEFSVLTTQN--- YAKFYYAKNKNGSSQVVYCFN  428
            |||| |||||||||||||||||||||||||||||    |||::|||||||||||||||||
Db    365  QKDGSQVENPNKEVAEPYSVEAYSDMQDSNYINPETFTPYGKFYYAKNKDKSSQVVYCFN  424

Qy    429  ADLKSPPDSEDGGKTMTPDFT-GEVKYTHIAGRDLFKYTVKPRDTDPDDFLKHIKKVIE  487
           ||||||||||||||||||||| |||||||:|||||:||:|||||||||||||| |||||
Db    425  ADLHSPPESEDGGGTIDPDISTMKEVKYTHTAGSDLFKYALRPRDTNPEDFLKHIKKVIE  484

Qy    488  KGYREKGQAIEYSGLTETQLRAATQLAIYFTDSAELDKDKL--- KDYHGFGDMNDSTL  543
           ||||||||: |||||||| |||||||||||| ||:|||||    |||||||||||| |:
Db    485  KGYNKKGDS--YNGLTETQFRAATQLAIYFTDTDLKTLKTYNNGKGYHGFESMDEKTL  542

Qy    544  AVAKILVEYAQDSNPPQLTDLDFFIPNNNKYQSLIGTQWHPEDLVDIIRMEDKK-EVIPV  602
           ||:|:|||||||||||| |||||||||||||||||||||||||||||||||||  |:|||
Db    543  AVTKELINYAQDSNSAPQLTNLDFFVPNNSKYQSLIGTEYHPDDLVAKDGKYSFNLKHG  602

Qy    603  THNLTLRKTVTGLAGDRTKDFHFEIELKNNKQELLSQTVKTDKINLEFKDGKATINLKHG  662
           ||||:|||| :|| | ||| ||||:|:: |||::|:| |||| :::|   ||| |||:||
Db    603  THSLTVKKTVVGELGDKITKGFQFELELKDKTGQPIVNTLKTNNQDLVAKDGKYSFNLKHG  662

Qy    663  ESLTLQGLPEGYSYLVKETDSEGYKVKVNSQEVANATVSKTGITISDETLAFENNKEPVVP  722
           :||||||||||||||||||||||||| || ||:|||:|||||||:||||||||||||| |
Db    663  DTIRIEGLPTGYSYTLKETEAKDYIVTVDNKVSQEAQSASENVTADRKVTFENRKDLVPP  722

Qy    723  TGVDQKINGYLALIVIAGISLGIW  746
           || ||||||||||||::| |:
Db    723  TGFITDGGTYLMLLLVPFGLLVW  746
```

RESULT 3

```
Q92B47
ID    Q92B47            PRELIMINARY;      PRT;    742 AA.
AC    Q92B47;
DT    01-MAY-1999 (TrEMBLrel. 10, Created)
DT    01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT    01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE    Cpa.
DE    CPA.
GN    Cpa.
OS    Streptococcus pyogenes.
OC    Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC    Streptococcus.
OX    NCBI_TaxID=1314;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=CS101;
RA    Podbielski A., Woischnik M., Leonard B.A.B., Schmidt K.H.;
RT    "Characterization of nra, a global negative regulator gene in group A
RT    streptococci.";
RL    Mol. Microbiol. 31:0-0(1999).
DR    EMBL; U49397; AAC97148.1; -.
SQ    SEQUENCE    742 AA;  83683 MW;  97A1FF44B4ECB944 CRC64;
```

```
Qy     17  RVLSKNSKRFTVTLVGVFLMIFPALVTSMVGAKTVEG----LVESSTPNAINPDSSSSEYRW  72
           || :: || :| :| ||||||||| ::|| ||::||    |:||:||||:|: ::| :|
Db     14  RVSNRKPKQLTVTLVGVFLMFLTLIVSSMRGAQSIFGEEKRIEEVSVPKIKSPDDA--YPW  71
```

RESULT 4

```
Q8K8U7
ID   Q8K8U7            PRELIMINARY;      PRT;   744 AA.
AC   Q8K8U7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Putative collagen binding protein.
GN   CBP OR SPYM3_0098.
OS   Streptococcus pyogenes (serotype M3).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=198466;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MGAS315 / Serotype M3;
RX   MEDLINE=22138808; PubMed=12122206;
RA   Beres S.B., Sylva G.L., Barbian K.D., Lei B., Hoff J.S.,
RA   Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA   Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA   Schlievert P.M., Musser J.M.;
RT   "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT   phage-encoded toxins, the high-virulence phenotype, and clone
RT   emergence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
DR   EMBL; AE014138; AAM78705.1; -.
KW   Collagen; Complete proteome; Hypothetical protein.
SQ   SEQUENCE 744 AA; 83650 MW; 38253EFA8DC7D4ED CRC64;
```

Query Match 49.8%; Score 1965.5; DB 2; Length 742;
Best Local Similarity 54.4%; Pred. No. 1.6e-100;
Matches 401; Conservative 116; Mismatches 195; Indels 25; Gaps 13;

```
Qy   20 SKNSKR--FTVTLVGVFLMIFALVTSMVGAKTVFGLVESSTPNAINPDSSSEYRWYGYES   77
         |:||:  |||  |||||||| ||| :| :|  |: : |||:| |||||||:|||||||  ::|
Db   11 SANNKRRQTTIGLLKVFLTFVALIGIVGFSIRAFGAEEQSVPN--RQSSIQDYPWYGYDS   68

Qy   78 YVRGHPYYKQFRVAHDLRVNLEGSRSRSYQYYCFNLKRAFFPLGSDSSSVKKWYKHDGISTKF  137
         : ||| ||:|: ||||||||:|:| ||| ||||||  :| ::||| ::||:||: |:::|||| :
Db   69 YPKGYPDYSPLKTYHNLKVNLEGSKDYQAYCFNLTKHFPSKSDSVRSQWYKLEGTNENF      128

Qy  138 EDYAMSPRITGDELNQKLRAVMYNGHPQNANGIMEGLEPLNAIRVTQEAVWYYSDNAPIS    197
          :|:| ||:| ||:||| || ::||| ||||||:|: ||||||:|||:||  :||:| || 
Db  129 IKLADKPRIEDGQLQQNILRILYNGYPNNRNGIMKGIDPLNAILVTQNAIW-YTDSAQI-   186

Qy  198 NPDESFKRESESNLVSTSQLSLMRQALKQLIDPNLATKMPKQVPDDFQLSIFESEDKGDK    257
         ||||||| |: ||:: || |||||:|||:||||||:|:||||:  ||| ||:||||| ||
Db  187 NPDESFKTEARSNGINDQOLGLMRKALKELIDPNLGSKYSNKTPSGYRLNVFESHD----  242

Qy  258 YNKGYQNLLSGGLVPFTKPPTPGDPPMPPNQPQTTSVLIRKYAIGDYSKLLEGATLQLTGD  317
          : :||||||:  |||:||| ||| | ||||| |||:||||||:|| ||||||||||:|  :
Db  243 --KPFQNLLSAEYVPDTPPKPGEE--PPAKTEKTSVIIRKYAEGD-SKLLEGATLKLSQI  297

Qy  318 NVNSFQARVFSSNDIGERIELSDGTYLTLELNSPAGYSIAEPITEKVEAGKVYTI-IDGK   376
         : :||||:|::|| |||:|||:|||||||::::||:||:||||  :||||||| ||:|
Db  298 EGSGFQEKDFQSNSLGETVELPNGTYLTLTETSSPDGYKIAEPIKFRVENKKVFIVQKDGS  357

Qy  377 QIENPNKEIVEPYSVEAYNDFEEFSVLT-TQNYAKFYYAKNKNGSSQVVYCFNADLKSPP   435
         |:|||:||: |||||||||||:::: |  |  : |||||:|:| :||||||||||:||||
Db  358 QVENPKEVAEPYSVEAYNDFMDEEVLSGFTPYGKFYYATNKDKSSQVVYCFNADLHSPP   417

Qy  436 DSEDGGKTMTPDFTT-GEVKYTHIAGRDLFKYTVKPRDTDPDTFLKHIKHKKVIEKGYREKG  494
         || | |:|| | ::| |||||| |||:|||||: ||||:|:|:|||||||||||||||::||
Db  418 DSYDSGETINPDTSTMKEVKYTHTAGSDLFKYALRPRDTNPEDFLKIHIKKVIEKGYKKKG   477

Qy  495 QAIEYSGLTETQLRAATQLAIYYFTDSAELDKDKL----KDYHGFGDMNDSTLAVAKILV   550
         ||:||:|||||:|||||||||||||||:||| |      ||||||| |::||||||| :||
Db  478 DS--YNLTETQFRAATQLAIYYFTDSADLKTLKTYNNGKGYHGFESMDEKTLAVTKELI    535

Qy  551 EVAQDSNPPQLTDLDFFIPNNNKYQSLIGTQWHPEDLVDIIRMEDKK-EVIPVTHNLTLR   609
         : || :: |||| ||| :|||::|:|||||: |||:|||:|||||||:|||||||:||:::
Db  536 TYAQNGSAPQLTNLDFPVPNNSKDQSLIGTECHPDDLVDVIRMEDKKQEVIPVTHSLTVK   595

Qy  610 KTVTGLAGDRTKDFHFEIELKNNKQELLSQTVKTDKTINLEFKDGKATINLKHGESLITLQG  669
         |||:| :||:||:| || ||||:  |:::| |||::||||:|||| : |||||| ::|:|:|
Db  596 KTVVGELGDKTKGFQFELELKDKTGQPIVNTLKTINNQDLVAKDGKYSFNLKHGDTIRIEG  655

Qy  670 LPEGYSYLVKETDSEGYKVKVNSQEVANATVSKTGTGTSDETLAFENNKEPVVPTGVDKOI  729
         || ||||:|| :::| ||::|:||::: |:|:|| || ||| ||||:|:: :|||| ||::
Db  656 LPTGYSYTLKEAEAKDYIVTVDNKVSOEAQSVGKDITEDKKYTFENRKDLVPPTGLLTDG   715

Qy  730 NGYLALIVIAGISLGIW                                             746
          ||||||:::| |:|:|
Db  716 AIYLWILLLVPLGLLVW                                             732
```

RESULT 5

```
Q8P2W3
ID   Q8P2W3            PRELIMINARY;      PRT;   524 AA.
```

Query Match 47.2%; Score 1862; DB 16; Length 744;
Best Local Similarity 50.2%; Pred. No. 8.6e-95;
Matches 371; Conservative 123; Mismatches 219; Indels 26; Gaps 11;

```
Qy   20 SKNSKR--FTVTLVGVFLMIFALVTSMVGAKTVFGLVESSTPNAINPDSSSEYRWYGYES   77
         |:||:  |||  |||||||| ||| :| :|  |: : |||:| |||||||:|||||||  ::|
Db   11 SANNKRRQTTIGLLKVFLTFVALIGIVGFSIRAFGAEEQSVPN--RQSSVQDYPWYGYDS   68

Qy   78 YVRGHPYYKQFRVAHDLRVNLEGSRSRSYQYYCFNLKRAFFPLGSDSSSVKKWYKHDGISTKF  137
         : ||| ||:|: ||||||||:|:| ||| ||||||  :| ::||| ::||:||: |:::|||| :
Db   69 YSKGYPDYSPLKTYHNLKVNLDGSKEYQAYCFNLTKHFPSKSDSVRSQWYKLEGTNENF     128

Qy  138 EDYAMSPRITGDELNQKLRAVMYNGHPQNANGIMEGLEPLNAIRVTQEAVWYYSDNAPIS    197
          :|:| ||:| ||:||| || ::||| ||||||:|: ||||||:|||:||  :||:| || 
Db  129 IKLADKPRIEDGQLQQNILRILYNGYPNDRNGIMKGIDPLNAILVTQNAIW-YTDSSYIS   188

Qy  198 NPDESFKRESESNLVSTSQLSLMRQALKQLIDPNLATKMPKQVPDDFQLSIFESEDKGDK    257
         ||||||| |: ||:: || |||||:|||:||||||:|:||||:  ||| ||:||||| ||
Db  189 NPDESFKSEGEKVELPNGTYVLSELKPQGYGVATPITFKVAAEKV--LIKNE        244

Qy  258 YNKGYQNLLSGGLVPFTKPPTPGDPPMPPNQPQTTSVLIRKYAIGDYSKLLEGATLQLTGD  317
         |: :||||||:  |||:||| ||| | ||||||
Db  245 --KTFQNLLSAEYVPDTPRKPGEE--PPAKTSVIIRKYAEGGYSKLLEGATLKLAQI     300

Qy  318 NVNSFQARVFSSNDIGERIELSDGTYLTLELNSPAGYSIAEPITEKVEAGKVYTII---D  374
Db  301 EGSGFQEKIFDSNKSGEKVELPNGTYVLSELKPQGYGVATPITFKVAAEKV--LIKNE    358

Qy  375 GKQIENPNKEIVEPYSVEAYNDFEEFSVLT-TQNYAKFYYAKNKNGSSQVVYCFNADLKS  433
Db  359 GQFVENQNKEIAEPYSVATPNDFEEIGYLSDFNNGKFYYAKNTNGTNQVVYCFNADLHS   418

Qy  434 PPDSEDGGKTMTPDFTT-GEVKYTHIAGRDLFKYTVKPRDTDPDTFLKHIKHKKVIEKGYRE  492
Db  419 PPDSYDHGANIDPDVSESKEIKYTHVSGYDLYKYAATPRDKDADFFLKIHKILDKGYKK   478

Qy  493 KGQAIEYSGLTETQLRAATQLAIYYFTDSAELDKDKL----KDYHGFGDMNDSTLAVAKI  548
Db  479 KGDT--YKILTEAQFRAATQLAIYYFTDSADLTTLKTYNDNKGYHGFDKLDDDATLAVVHE 536

Qy  549 LVEVAQDSNPPQLTDLDFFIPNNNKYQSLIGTQWHPEDLVDIIRMEDKK-EVIPVTHNLT  607
Db  537 LITYAEDVLPMTQNLDFFVPNSSRYQALIGTQYHPNELIDVISMEDKHQAPIIPITHKLT  596

Qy  608 LRKTIVTGLAGDRTKDFHFEIELKNNKQELLHLKSSDGQAISGTYPTNSGELTVTDGKATFTLKDGESLIV  667
Db  597 ISKTVTGTIADKKKEFNFELHLKSSDGQAISGTYPTNSGELTVTDGKATFTLKDGESLIV         656

Qy  668 QGLPEGYSYLVKETDSEGYKVKVNSQEVANATVSKTGTGTYPTNSGELTVTDGKATFTLKDGESLIVDQ  727
Db  657 EGLPSGYSYTEITGASDYEVSVSNGKNAPDGKATKASVKEDETVAFENRKDLVPPTGLLTT     716

Qy  728 KINGYLALIVIAGISLGIW                                           746
Db  717 DGAIYLWILLLVPFGLLVW                                           735
```

```
Q8P2W3
ID   Q8P2W3            PRELIMINARY;      PRT;   524 AA.
```

```
AC   Q8P2W3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Putative collagen binding protein)
GN   CBP OR SPYM18_0126.
OS   Streptococcus pyogenes (serotype M18).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=186103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MGAS8232 / Serotype M18;
RX   MEDLINE=21927593; PubMed=11917108;
RA   Smoot J.C., Barbian K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA   Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA   Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA   Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT   "Genome sequence and comparative microarray analysis of serotype M18
RT   group A Streptococcus strains associated with acute rheumatic fever
RT   outbreaks.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
KW   Collagen; Complete proteome; Hypothetical protein.
SQ   SEQUENCE  524 AA;  57872 MW;  9AA79F7C1A6D6DD0 CRC64;

Query Match             31.8%;  Score 1254.5;  DB 16;  Length 524;
Best Local Similarity   54.2%;  Pred. No. 2e-61;
Matches  254;  Conservative  69;  Mismatches  129;  Indels  17;  Gaps  9;

Qy   287 QPQTPSVLIRKYAIGDYSKLLEGATIQLTGDNVNSFQARVFSSNDIGERIELSDGTYTLT 346
             |:|||||||||||||||||||||||:|||||||||:|||||||:|||||||||||||||
Db    52 ETKTKTSVIRKYAEGDYSKLLEGATLKLAQIEGSGFQEQSFESSTSGQKLQLSDGTYILT 111

Qy   347 ELNSPAGYSIAEPITFKVEAGKVYTI-IDGKQIENPNKEIVEPYSVEAYNDFEEFSVL-- 403
             ||||||||||||||||||||||||||  ||| || ||||:||||||||||:|||||||
Db   112 ETKSPQGYETAEPITFKVTAGKVFIKGKDGQFVENQNKEVAEPYSVTAYNDFDDSGFINP 171

Qy   404 -TTQNYAKFYYAKNKNGSSQVYYCFNADLKSPPDSEDGGKTMTPDFTTG-EVKYTHIAGR 461
              :||:||||||||||||||||||||||::|||:|||||||||||||| ||||:|||||
Db   172 KTFTFPYGKFYYAKNANGTSQVYYCFNVDLHSPPDSLDKGEIDPDFNEGKEIKYTHILGA 231

Qy   462 DLFKYTVKPRDTDPDTFLKHIKKVIEKGYREKGQAIEYSGLTETQLRAATQLAIYYFTDS 521
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   232 DLFSYANNPRASTNDELLSQVKKVLEKGYRD--DSTTYANITSVEFRAATQLAIYYFTDS 289

Qy   522 AELDKDKLKDYHGFGQDMNDSTLAVAKILVEYAQD-SNPPQLTDLDFFIPNNNKYQSLIGT 580
             :||||||||||||||||:||||||||||:||||||||    ||:|||||||||||||||
Db   290 VDL--DNLADYHGFGQALTTEALNATKEIVAYAEDRANLPNISNLDFYVPNSNKYQSLIGT 347

Qy   581 QWHPEDLVDIIRMEDKK-EVIPVTHNLTRKTVTGLAGDRTKDFHFEIELKNNKQELLSQ 639
             |||:|||||:|||  |:||||:|||||:|:||||||:||| |||:||||||||||||:
Db   348 QYHPESLVDIIRMEDKQAPIPITHKLTISKTVTGTIADKKEFNPEIHLKSSDGQAISG 407

Qy   640 TVKIDKTNLEFFKDGKATINLKHGESLTLQGLPEGYSYLVKEETDSEGYKVKVNSQEVANAT 699
             |||:||||||:|||||||||||||||:|||||||||||||||||||||||||||||| |
Db   408 TYPINSGELTVTDGKATFTLKDGESLIVEGLPSGYSYEITETGASDYEVSVNGKNAPDGK 467

Qy   700 VSKTGTITSDETLAFENNKEPVVPTGV--DQKINGYLALIVAGISLGIW 746
             :||||  |||:|  |||||     |    || |||||||||| ||| |
Db   468 ATKASVKEDETITFENRKDLVPPTGITTDGAIYLWILLLLVL----LGLW 512
```

```
RESULT 6
Q54953

ID   Q54953      PRELIMINARY;      PRT;      659 AA.
AC   Q54953;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Protein F.
GN   PRTF.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JRS75;
RX   MEDLINE=95020565; PubMed=7934855;
RA   Sela S., Aviv A., Burstien I., Tovi A., Caparon M.G., Hanski E.;
RT   "Protein F: An adhesin of Streptococcus pyogenes binds fibronectin via
RT   two distinct domains.";
RL   Mol. Microbiol. 10:1049-1055(1993).
DR   EMBL; L10919; AAA26964.1; -.
DR   InterPro; IPR004237; Fn_bind.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR006192; LPXTG.
DR   Pfam; PF02986; Fn_bind; 6.
DR   TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR   PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
SQ   SEQUENCE  659 AA;  73628 MW;  D0CEB91D159726BD CRC64;

Query Match             16.1%;  Score 633.5;  DB 2;  Length 659;
Best Local Similarity   26.1%;  Pred. No. 5.8e-27;
Matches  203;  Conservative  128;  Mismatches  237;  Indels  211;  Gaps  30;

Qy    20 SKNSKRFTVTLVGVFLMIFALVTSMVGAKTVFGLVESSTPNAINPDSSSEYRWYGYESYV 79
             ||:||  ||:||| |||| ||  | | || :||  ||:||||  ||||:||:| |||:
Db    24 TKRKRRFAVTLLGVGFFMLLACAGAIGFGQVAYAADEKTVPSHSSP--NPEFPWYGYDAYG 81

Qy    80 RGHPYYKQFRVAHDLRVNLEGSRSYQYCFNLKKAFPLGSDSSVKKWYKTKHDGISTKFED 139
             |||:||   |  | ||||| | |||||||| |     | | ||:|   | | | || |:
Db    82 KEYPGYNIWTRYHIDLRVNLNGSRSYQYCFNIQSNYPSQKINGFIKNWFKKIEGNGKSFVD 141

Qy   140 YAMSPRITGDELNQKLRAVMYNGHPQNANGIMEGLEPLNAIRVTQERAVWYSDNAPISNP 199
             ||:||:|||:|||||||| | | ||||| |:| |: | || ||| |:| | |: |:||
Db   142 YAHTTKLGKEELPQRLLSLLYNAYPNDANGYMKGLEHNAIIVTQYAVWHYSDNSQYQF- 200

Qy   200 DESFKRESESNLVSTSQLSLMRQALKQLIDPNIATKMPKQVPDDFQLSIFESEDKGDKYN 259
             ||| ||||  ||  | ||:| | ||||| ||| |  ||:|| ||||| |:|||||:|||
Db   201 ETLWESEAKEGKISRSQVTLMREALKLIDPNLEATAVNKIPSGYRLNIFESE------N 254

Qy   260 KGYQNLLSGGLVPTKPPTPGDPPMPNQPQOTTPSVLIRKYAIGDYSKLLEGATL----N 315
             |||:|||| |||||  | |||| |||||| ||||||||||||||||||||||||    :
Db   255 EAYQNLLSAEYVPDDPPKPGE--TSEHNPKTPB----------------LDGTPIPEDPKHP 298

Qy   316 GDNVNSFQARVFSSNDIGERIELSDGTYTLTELNSPAGYSIAEPITFKVEAGKVYTIIDG 375
             | || |||||||||||||||||||||||| |||||||| | ||:||||| ||||| |
Db   299 DDNLEPILPPVMLD----GEEV----------PEVPSESLEPALPFLMPE-----------LDG 337

Qy   376 KQIENPNKEIVE-PYSVEAYNDFEEFSVLTIQNYAKFYYAKNKNGSSQVYYCFNADLKSP 434
             |||||| ||| | |||||| |||||| || || || |||||||||| ||||||||| |||
Db   338 QEV--PEKPSIDLPIEVPRYE----------------------------------FNNKDQSP 364

Qy   435 PDSEDGGKTMTPDFTTGEVKY-THIAGRDLFKYTVKPRDTDPDTFLKHIKKVIEKGYREK 493
             |||| ||||||||| | | |  || || |||| | || ||||||||||||||||||||||
Db   365 LAGE------------------------SGETEYITEVIGNQ--------QNPVDIDK------- 399

Qy   494 GQAIEYSGLTETQLRAATQLAIYYFTDSAELDKDKLKDYHG------------------ 536
             | |||||||||||||||||||||||||| ::    :
Db   400 GNMVETEDTKEPEVLMGGQ----------SESVEFFTKDTQTGMSGQTTPQVETEDTKEPEVLMG 453

Qy   537 DMNDSTLAVAKILVEYAQDS----------NPPQLTDLDFFIPNNNKYQSLIGTQWHPEDLVD 589
             ::|                 |::||||||||||||||||||||:
Db   454 GQSES---------VEFFKDTQTGMSGQTTPQIETEDTKEPE------- 486

Qy   590 IIRMEDKKEVIPVTHNLTRKTVTGLAGDRTKDFHFEIELKNNKQ-ELL----SQTVKTD 644
             ||||||| ||| ||||  | |||||||| ||     | |:|
Db   487 -VLMGGQSESVEFT-----KDTQTGMSGQTTP----QVETEDTKEPEVLMGGQSESVEFT 536

Qy   645 KTNLEFFKDGKATINLKHGESLTLQGLPEGYSYLVKEETDSEGYKVKVNSQEVANATVSKTG 704
             | ||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db   537 KDTQTGMSGQTTPQIETEDTKEPEVLMGGQSESVEFT----------------------KDTQTGMSGFSETA 587

Qy   705 ITTSDET------LAFENNK------------------------------EPVVP-TGVDQKINGYLALIVIAGISL 743
             :|| |        ||| || |
Db   537 KDTQTGMSGQTTPQIETEDTKEPEVLMGGQSESVEFT 587
```

```
Db  588 TVVEDTRPKLVFHFDNNEPKVEENREKPTKNITPILPATGDIENVLAFIGILILSVLSI 646
        ||||| || :||||||||:| |||| ||  : |||||||||||||||||||:||||||||
```

RESULT 7

```
ID   Q8RP54      PRELIMINARY;      PRT;      698 AA.
AC   Q8RP54;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Protein F.
GN   PRTF1.12.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bessen D.E., Kalia A.;
RX   MEDLINE=21843101; PubMed=11854196;
RC   STRAIN=A735;
RT   "Genomic localization of a T serotype locus to a recombinatorial zone
RT   encoding extracellular matrix-binding proteins in Streptococcus
RT   pyogenes.";
RL   Infect. Immun. 70:1159-1167(2002).
DR   EMBL; AF447492; AAL86405.1; -.
DR   InterPro; IPR004237; Fn_bind.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR006192; LPXTG.
DR   Pfam; PF02986; Fn_bind; 5.
DR   PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
DR   TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
SQ   SEQUENCE   698 AA;  77312 MW;  F73A41885D273952 CRC64;

Query Match             16.0%;  Score 633;  DB 2;  Length 698;
Best Local Similarity   25.4%;
Matches  207;  Conservative  129;  Mismatches  235;  Indels  244;  Gaps  31;

Qy   18 VLSKNSRFTVTLVGVFLMIFALVTSMVGAKTVFGLVESSTPNAINPDSSSEYRWYGYES 77
         |:|| :|||||| |||||||| ||||  |:||||||:|||||||||||||||||:|| :
Db   27 VHTKRKRFAVTLVGVFLLLACAGAIGFGQVAYAADEKTVPNFKSPD--PDYPWYGYDS 84

Qy   78 YVRGHPYYKQFRVAHDLRVNLEGSRSYQVYCFNLKKAFPLGSDSSVKKWYKKHDGISTKF 137
         | ||     | ||:|||||||:||:|||:||||| | |:   :  ::  |||:|||  ||
Db   85 Y-RG-----IFARYHNLKVNLKGSKEYQAYCFNLTKYFPRPIVYSTTNNFYKKIDGSGSAF 138

Qy  138 EDYAMSPRITG---DELNQKLRAVMYNGHPQNANGIMEGLEPLNAIRVTQEAVWYYSDNA 194
         : ||  |    :: |:|:| : :|: ||  :|||| |:|||:||||:|||:|:||||||:|
Db  139 KSYAANPRVLDENLDKLEKNILNVIYGNGYKSNANGFMNGIEDLNAILVTQNAIWYYSDSA 198

Qy  195 PISNPDESFKRESESNLVSTSQLSLMRQALKQLIIDPNLATKMPKQVPDDFQLSIFESEDK 254
         | :: ::  :RE   |:S SQ::LMR:ALK:LIIDPNL   :    ||   :L ||::::
Db  199 PLNDVNKMWEREVRNGEISESQVTLMREALKKLIIDPNLEATAAANKIPSGYRLNFKSE-- 256

Qy  255 GDKYNKGYQNLLSGGLVPTKPPTPGD------------------PPMPPN 286
             :::YQNLL:   |P  PP PGD                  P  P
Db  257 ----NEDYQNLLSAEYVPDDPPKPGDTSEHNPKTPELDGTPIPEDPKRPDESSEPALPPL 312

Qy  287 QPQTTSVLIRKYAIGDYSKLLEGATLQLTGDNYNSFQQARVFSSNDIGERIELS------- 339
          P:  ::           :L E   L L GD  :                 :EL
Db  313 MPELDGEEVPEVPSESLEPALPPLMPELDGEEVPEVPS-------- 363

Qy  340 -DGTYLTELNSPAGYSIAEPITFKVEAGKVYTI----------- 383
          DG  :         S  EP      : G:V
Db  364 LDGE--EVPEVPSESLEPALPPLMPELDGEEV--PEK 420

Qy  384 EIVE-PYSVEAVNDFEEFSVLTQNYAKFYAKNKNGSSQVVYCFNADLKSPPDSEDGGK 442
          :V: P: VE V            :       ::  S      N DL SP   E
Db  421 PSVDLPIEVPRVE----------------------------FNNKDQSPLAGE--- 445

Qy  443 TMTPDFTTGEVKY-THIAGRDLFKYTVKPRDTDPDTFLKHIKVIEKGYREKCQAIEYSG 501
                  GE        G:           :   :    K    : G  E:C  :EY G
Db  446 ------------------SGETEYITEVYGNQ-----QNPVDIK------KLPNETGF--SGNNVETED 484

Qy  502 LTETQLRAATQLAIYYFTDSAELDKDRLKIDYHGFGDMNDSTLAVAKILVEYAQDSNPPQL 561
Db  485 TKEPEVLMGGQ-------SESVEFTKDTQTGMSG------QTTPQVE-----TEDIKEPEV 527

Qy  562 TDLDFFIPNNNKYQSLLIGTQWHPEDLVDIIRMEDKKEVIPVTHNLTLRKTVTGLAGDRTK 621
Db  528 ------------LMGGQ--------SESVEFT-----KDTQTGMSGQTTP 552

Qy  622 DFHFEIELKNNKQELL----SQTVKTDKTNLEFKQDGKATINLKHGESLTLQGLPECYSY 676
Db  553 ----QVETEDTKEPGVLMGGQSESVEFTKDTQTGMSGQTTPQVE----- 592

Qy  677 LVKEFDSEGYKVKVNSQEVANATVSKTGITS-DETLA-----------FENNK----- 717
Db  593 -TEDTKEPGVLMGGQSESVEFTKDTQTGMSGGFSETVTIVEDTRPKLVFHFDNNEPKVEEN 651

Qy  718 EPVVP-TGVDQKINGYLALIVIAGISL 743
Db  652 REKPTKNITPILPATGDIENVLAFLGILILSVLSI 686
```

RESULT 8

```
ID   O33709      PRELIMINARY;      PRT;      685 AA.
AC   O33709;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Fibronectin/fibrinogen binding protein F.
GN   PRTF15.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98129085; PubMed=9467904;
RA   Katerov V., Andreev A., Schalen C., Totolian A.A.;
RT   "Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
RT   also binds human fibrinogen: isolation of the protein and mapping of
RT   the binding region.";
RL   Microbiology 144:119-126(1998).
DR   EMBL; AF009908; AAC38155.1; -.
DR   InterPro; IPR004237; Fn_bind.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR006192; LPXTG.
DR   Pfam; PF02986; Fn_bind; 5.
DR   PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
DR   TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
SQ   SEQUENCE   685 AA;  75988 MW;  9F4633A35C5D4DD3 CRC64;

Query Match             15.5%;  Score 611;  DB 2;  Length 685;
Best Local Similarity   25.9%;  Pred. No. 1.1e-25;
Matches  210;  Conservative  117;  Mismatches  237;  Indels  246;  Gaps  32;

Qy   20 SKNSRKFTVTLVGVFELMIFALVTSMVGAKTVFGLVESSTPNAINPDSSSEYRWYGYESYV 79
          ::|| ||||||||| :|:| ||| :|:|  |:|||||:|||||||:| |  || |||:||
Db   24 TKRKRRFAVTLVGVFFMLFRACAGSIGFGQVAYAADEKTEPNRIN--SNPEPWYGYDSYS 81

Qy   80 RGHPYYKQFRVAHDLRVNLEGSRSYQVYCFNLKKAFPLGSDSSVKKWYKKHDGISTKFED 139
         :|        |||:|||||:||  :|  :|||:|:| |  ||:| |||||:|||:| :|
Db   82 GRLLRY-------HNLKVN--GSKEYQAYCFNLKRFEPKKEESSSPNWYKKLDGSTEFKK 133

Qy  140 YAMSPRITGDELNQKLRAVMYNGHPQNANGIMEGLEPLNAIRVTQEAVWYYSDNAPISNP 199
         ||  :| :||:|: :|: :|||  |:::|:| :|::PLNAI:VTQ:A:WYYSD:API :
Db  134 YAENPFSGEELRHILKVLYNGYP-NSNEIMKGIDPLNAILVTQNAIWYYSDSAPINDI 192

Qy  200 DESFKRESESNLVSTSQLSLMRQALKQLI--DPNLATKMPKQVPDDFQLSIFESEDKGDK 257
         :  ::RESESNL:  :QL:LMR:AL::LI  DPNL    :   ||  ::||FES D
Db  193 NFFTSEANDLNIPPQQLTLMREALRKLISSDENLV----KQVPSNFKLSIFESSD---- 244
```

```
RESULT 9
ID   Q01924      PRELIMINARY;   PRT;   638 AA.
AC   Q01924;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Fibronectin-binding protein precursor.
GN   SFB.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE OF 1-303 FROM N.A.
RC   STRAIN=DSM 2071;
RX   MEDLINE=92363585; PubMed=1386839;
RA   Talay S.R., Valentin-Weigand P., Jerlstrom P.G., Timmis K.N.,
RA   Chhatwal G.S.;
RT   "Fibronectin-binding protein of Streptococcus pyogenes: sequence of
RT   the binding domain involved in adherence of streptococci to epithelial
RT   cells.";
RL   Infect. Immun. 60:3837-3844(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 2071;
RX   MEDLINE=95089690; PubMed=7527894;
RA   Talay S.R., Valentin-Weigand P., Timmis K.N., Chhatwal G.S.;
RT   "Domain structure and conserved epitopes of Sfb protein, the
RT   fibronectin-binding adhesin of streptococcus pyogenes.";
RL   Mol. Microbiol. 13:531-539(1994).
```

```
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 2071;
RA   Talay S.R.;
RL   Submitted (JUL-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X67947; CAA48133.1; -.
DR   InterPro; IPR004237; Fn_bind.
DR   Pfam; PF02986; Fn_bind; 5.
KW   Signal.
FT   SIGNAL        1     49       POTENTIAL.
SQ   SEQUENCE   638 AA;  71059 MW;  D55E2A4E5708F3E0 CRC64;

Query Match             15.3%;  Score 603;  DB 2;  Length 638;
Best Local Similarity   27.4%;  Pred. No. 2.7e-25;
Matches  207;  Conservative  118;  Mismatches  248;  Indels  182;  Gaps  32;
```

Left alignment (RESULT 9):

```
Qy   258 YNKGYQNLLSGGLVPFTKPPTPGD-----------------------PPMPENQPQ 289
Db   245 --KSYQNLLSAEYVPDDPPKPGDTSEHNPKTPELDGTPIPEDPKRPDESSGPALPPLMPE 302

Qy   290 TTSVLIRKYAIGDYSKLLEGATLQLTGDVNVNSFQARVFSSNDIGERIELS--------DG 341
Db   303 LDGEEVPEVPSESLEPALPPLVPELDGEEVPEVPS----------ESLEPALPPLMPELDG 353

Qy   342 TYTLTELNSPAGYSIAEPITFKVEAGKVYTI--------------IDGKQIENPNKEIV 386
Db   354 -QEVPEVPSESLEPALPPLMPELDGEEIPEVPSESLEPALPPLMPELDGQEV--PEKPSI 410

Qy   387 E-PYSVEAYNDFEEFSVLTTQNYAKFYYAKNKNGSSQVVYCFNADLKSPPPDSEDGKTMT 445
Db   411 DLPIEVPRYE----------------------------FNNKDQSPLAGE---------- 432

Qy   446 PDFTTGEVKYTHIAGRDLFKYTVKPRDTDPDTFLKIHIKKVIEKGYREKGQAIEYSGLTET 505
Db   433 ----SGETDYI----IEVYGNQQNPVDIDK----KLPNETGF--SGNMVETEETKEP 475

Qy   506 QLRAATQLAIYYFTDSAELDKDKLKDYHG----------FGDMNDSTLAVAKI 548
Db   476 GVLMGGQ------SESVEFTKDTQAGMSGQTTPQVETEETKEPGVLMGGQSES------- 522

Qy   549 LVEYAQDSNPPQLTLDLDFFIPNNNKYQSLIGTQWHPEDLVDLIIRMEDKKEVIPVTHNLTL 608
Db   523 -VEFTKDT--------QTGMSGQTAPQ----VETEDTKE-PEV---- 551

Qy   609 RKTVIGTGLAGDRTKDFHFEIELKNNKQELLSQTVKTDKTNLEFKDGKATINLKHGESLTLQ 668
Db   552 -----LMGGQSGSVEFK----KDTQAGMSGQTA----PQVETEDTKEPEVLMGGQSESSVE 598

Qy   669 GLPEGYSYLVKETDSEGYKVKVNSQEVANATVSKTGITSDETLAFENNK---------- 717
Db   599 FTKD--------TQTGSSGFSETVSIVEDTRPKL--------VFHFDNNEPKVEEHREKPT 643

Qy   718 ---EPVVP-TGVDQKINGYLALIVIAGISL 743
Db   644 KNITPILPATGDIENVLAFLGILILSVLSI 673
```

Right alignment:

```
Qy    20 SKNSKRFTVTLVGVFLMIFALVTSMVGARTVFGLVESSTPNAINPDSSSEYRWYGYESYV 79
Db    17 TKRRRFAVTLVGVFFMILASAGAIGFGQVAYAADEKTVPHRVS--QNPEFPWYGYDFY- 73

Qy    80 RGHPYYKQFRVAHDLRVNLEGSRSYQYCFNLKKAFPLGSDSSVKKWKKHDGISTKFPD 139
Db    74 KG-PYTRY----HNLQLNLNGSKTYQAVCFNLKFEPKKEGSYFPNWVKRWDGSEETVK 128

Qy   140 YAMSP-------RITGDELNQKLRAVMYNGHPQNANGIMEGLEPLNAIRVTQEAVWYSD 192
Db   129 YADNPRKDNESSRVIDVELEKNILRVLYINGYPNNGNGIMEGLEPLNAILVTQNAVWYYSD 188

Qy   193 NAPISNPDESFKRESESNLVSTSQLSLMRQALKQLIDPNLNATRMPKQVPDDFQLSIEFSE 252
Db   189 NSSIFNTDNFFTEAKIDLNIKPEQLSLMRVALKKLIDPKLSEESLKPVSTFRLNIFESQ 248

Qy   253 DKGDKYNKGYQNLLSGGLVPFTKPPTPGDPPMPPNQPQTTSVLIRKYAIGDYSKLLEGATL 312
Db   249 D------KLYQNLLSAEVPENNPKPGE--TPEHGPKIPEL---DGTPIPEGP-- 290

Qy   313 QLTGDVNVNSFQARVFSSNDIGERIELSDGTYTLTELNSPAGYSIAEPITFKVEAGKVYTI 372
Db   291 QRPNESLEPTLPPV----------MLDG-QEVPEVPSESLEPALPPLMPELDGQEVPEV 338

Qy   373 --------IDGKQI-ENPNKE--IVEPYSVEAYNDFEEFSVLTTQNYAKFYYA 414
Db   339 PSESLEPALPPLMPELDGQEVPEKPSVDLPIEDPR--------------- 373

Qy   415 KNKNGSSQVVYCFNADLKSPPPDSEDGKTMTPDFTTGEVKY-THIAGRDLFKYTVKPRDT 473
Db   374 --------YEPNNKDQSPLAGE--------------SGETEYITEVYGNQ---- 407

Qy   474 DPDTFLKIHIKKVIEKGYREKGQAIEYSGLTETQLRAATQLAIYYFTDSAELDKDKLKDYH 533
Db   408 DK-------KLPNETGF--SGNMVETEDTKEPGVLMGGQ----- 452

Qy   534 GFGDMNDSTLAVAKILVEYAQDSNPP------QLTDLDFFIPNNNKYQSLIGTQWHPEDL 587
Db   453 G------QTPQVE--------TEDTKEPGVLMGGQSESVEFTKDTQTGMSGQTASQVETEDT 501

Qy   588 VDI-IRMEDKKEVIPVTHNLTLRKTVIGLAGDRTKDFHFEIELKNNKQELLSQTVKTDKT 646
Db   502 KEPGVLMGGQSESVEFT----KDTQTGMSGQTTP----QVETEDTKEPGVLMGGQSE-- 550

Qy   647 NLEF-KDGKATINLKHGESLTLQGLPEGYSYLVKETDSEGYKVKVNSQEVANATVSKTGI 705
Db   551 SVEFTKDTQTGMS---GFSETVTIVEDTRPKLVFHFDNNEPKVEENREKPT------ 598

Qy   706 TSDETLAFENNKEPVVP-TGVDQKINGYLALIVIA 739
Db   599 ------KNITPILPATGDIENVLAFLGILILS 624
```

```
RESULT 10
ID   Q47942      PRELIMINARY;   PRT;   580 AA.
AC   Q47942;
```

```
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Group G streptococcal fibronectin binding protein.
GN   GFBA.
OS   Streptococcus sp. (Lancefield group G).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1750;
RX   MEDLINE=96239026; PubMed=8675316;
RA   Kline J.B., Xu S., Bisno A.L., Collins C.M.;
RT   "Identification of a fibronectin-binding protein (GfbA) in pathogenic
RT   group G streptococci.";
RL   Infect. Immun. 64:2122-2129(1996).
DR   EMBL; U31115; AAB06623.1; -.
DR   InterPro; IPR004237; Fn_bind.
DR   InterPro; IPR001890; Gram_pos_anchor.
DR   Pfam; PF02986; Fn_bind; 4.
DR   TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
SQ   SEQUENCE   580 AA;  64975 MW;  6994BF41C573A70 CRC64;

Query Match             13.3%;  Score 523;  DB 2;  Length 580;
Best Local Similarity   24.8%;  Pred. No. 6.2e-21;
Matches  180;  Conservative 106;  Mismatches 250;  Indels 190;  Gaps  23;

Qy    20 SKNSKRFTVTLVGVFLMIFALVTSMVGAKTVFGLVESSTPNAINPDSSSEYRWYGYESYV  79
         |:|  |||||||| : ||||||||:|: ||||  |  ::||||  :  ||||||||:|
Db    23 TKRKRRFAVTLVGVFPMLLACAGAIGFGQVAYAADEKTVPNFKSPD----PDYPWYGIDAYT  80

Qy    80 RGHPYYKQFRVAHDLRVNLEGSRSYQVYCFNLKKAAFPLGSDSSVKKWYKKHDGISTKFED  139
         ||| |||||||  |||: |||||   ||::|||:||||||||||  ||||||:|  ||||
Db    81 GAFLRY------HDLNVNLEGSTPYQVYCFNLVRQEPSKVNGFRKFRFKKVDGDNAVFKQ  134

Qy   140 YAMSPRITGDELNQKLRAVMYNGHPQNANGIMEGLEPLNAIRVTQEAVWYISDNAPISNP  199
         ||||||||||||| | : :||   ::||||||| |:|||||||:: ||| |:||||: |:
Db   135 YAANPRYIDGDLERNILNVIYNGYPSDANGIMKGLDRYNAILVTQTAIWYSDNSKFI-V  193

Qy   200 DESFKRESESNLVSTSQLSLMRQALKQLIDPNLATKMPKQVPDDFQLSIFESEDKGDKYN  259
         | :|| ||| |:|: | | |:|||:| ||::   |:|::|| || ::| :| |||:| |
Db   194 DEQWKSELRDGEISESQVKLMRDALRKLISPDLEKTVKNQLPENYKLNLFISKD------  247

Qy   260 KGYQNLLSGGLVPTKPPTPGDPDPMPNQPQOTTSVLIRKYAIGDYSKLLEGATLQLTGDNV  319
         || |||  ||||||| ||| |  :|| |:|:|  ::|:| ||:| ||:|||:|  |:
Db   248 NSIQNLLSTEYVQDNLQKPGEEP---------------------------------  270

Qy   320 NSFQARVFSSNDIGERIELSDGTYTLTELNSPAGYSIAEPITEKVERAGKVVTIIDGKQIE  379
         ||  | |:||| ||:| | | | :|| |:  |  |: |::||  | ||: |:|:|||||:|
Db   271 ----KEESPKEENPKIPDLDG-HEIPEVPQEPSDPVIPPV--------------ILDGEEVP  313

Qy   380 NPNKEIVEPYSVEAVNDFEEFSVLTIQNYAKFYYAKNKNGSSQVV-----YCFNADLKSPP  435
         || |:|||  :|  ||: ||| |||:|  ||| :|||:| |:| |     |||:|  ||
Db   314 EVPSESLEPALPLMPELDGQEV----------------PEKPSIDLPIEVPYRYENNKDQSPL  361

Qy   436 DSEDGGKTMTDPDTTGEVKY-THIAGRDLFKYTVKPRDTDPDTDFLKIHKKVIEKGYREKG  494
         | :|  |:|:|| :||   : ||:|||||||| ::| : :|||:|| ||:|:  |||||:
Db   362 AGE-------------------SGEYEYITEVYGNO---------QNPVDIDK----KLPNETGF--SG  396

Qy   495 QAIEYSGLTETQLRAATQLAIYYFTDSAELDKDKLKDYHGFGDMNDSTLAVAKILVEYAQ  554
         | |:| | |:||||  ||||: || |||: ||:|    :|   |:||| ||:|||:|:|
Db   397 NMVETEDTKEPEVLMGGO--------SESVEFTKDTQTGMSG------------------------  430

Qy   555 DSNPQQLTLDFFIPNNNKYQSLIGTQWHPEDLVDIIRMEDKKEVIPVTHNLTLRKTVTG  614
         |||||:||||||||||||:|:|||| | |||:|||:|||| |:|||:|| |||:|||
Db   431 -QTTPQVETEDTKEPG-----------------------------------VLMGQSESVEFT----  464

Qy   615 LAGDRTKDFHFIELKNNKQELLSQTVKTDKTNLEFKDGKATINLKHGESLILQGLPEGY  674
         | |||||||| | |:||:|:|| | ||:|| |||  ||| |:||  :|:| ||  | |
Db   465 MSGQTTP---QVETEDTKEPGVLMGGQSESVGFT-KDTQTGMS----GFSETATVVEDTR  516

Qy   675 SYLVKEFDSEGYKVKVNSQEVANATVSKTGITSDETLAFENNKEPVVP-TGVDQKINGYL  733
```

```
Db   517 PKLVFHFDNNEPKVEENQEKPT-----------------------------------KNITPILPATGDIENVLAFL  558
         ||| |  | |:| ::: ::                                        | :: | ||
Qy   734 ALIVIA  739

Qy   559 GILLS  564
```

```
RESULT 11
ID   O33715      PRELIMINARY;      PRT;   106 AA.
AC   O33715;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   Fibronectin binding protein F (Fragment).
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M4 strain Umea;
RX   MEDLINE=98129085; PubMed=9467904;
RA   Katerov V., Andreev A., Schalen C., Totolian A.A.;
RT   "Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
RT   also binds human fibrinogen: isolation of the protein and mapping of
RT   the binding region.";
RL   Microbiology 144:119-126(1998).
DR   EMBL; AF009914; AAC38161.1; -.
FT   NON_TER       1      1
FT   NON_TER     106    106
SQ   SEQUENCE   106 AA;  12329 MW;  58FDF55C1A9138E2 CRC64;

Query Match             6.9%;  Score 272.5;  DB 2;  Length 106;
Best Local Similarity  49.1%;  Pred. No. 3.9e-08;
Matches   53;  Conservative  17;  Mismatches  31;  Indels  7;  Gaps  2;

Qy    51 FGLV-----ESSTPNAINPDSSSEYRWYGYESYVRGHPYYKQFRVAHDLRVNLEGSRSYQ  105
         ||:|     ||:||||:|||||||||||||||||||||:||||||||||||||||||||
Db     1 FGQVAYSADEKTVPNFKSPD--PDYPWYGIYAYGKGYPGYDISKYYHDLRVNLNGSQVYQ  58

Qy   106 VYCFNLKKAAFPLGSDSSVKKWYKKHDGISTKFEDYAMSPRITGDELNQ  153
         |||:| |||  :|:||||||||||||||||||||||||||||||||||
Db    59 VYCFNIQKIPPYNVKSVTQKWFKKAEGNSDTFGPYAMNPRVQGEELSQ  106
```

```
RESULT 12
ID   O33714      PRELIMINARY;      PRT;   104 AA.
AC   O33714;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   "Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
DE   also binds human fibrinogen.";
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M4 strain 281C;
RX   MEDLINE=98129085; PubMed=9467904;
RA   Katerov V., Andreev A., Schalen C., Totolian A.A.;
RT   "Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
RT   also binds human fibrinogen: isolation of the protein and mapping of
RT   the binding region.";
RL   Microbiology 144:119-126(1998).
DR   EMBL; AF009913; AAC38160.1; -.
FT   NON_TER       1      1
FT   NON_TER     104    104
SQ   SEQUENCE   104 AA;  12055 MW;  CF6C1A9E05A72850 CRC64;
```

```
Qy      51 FGLV-----ESSTPNAINPDDSSSEYRWYGYESYVRGHPYYKQFRVAHDLRVNLEGSRSYQ 105
              ||| |    |:|| ::||| :| ||:|  | | |:|||  | |||| :|||||||||| ||
Db       1 FGQVAYSADEKTVPNFKSPD--PDIPWYGYDAYGKGYPGYPGTDISKYYHDLRVNLNGSQVYQ 58

Qy     106 VYCFNLKKAFPLGSDSSSVKKWYKHDGISTKFEDYAMSPRITGDEL 151
              ||| || : :| :::|| ::|||||| | ||||| ||:|| ||| |
Db      59 VYCFNIQKIIPPYNVKSVTQKWFKAEGNSDTFGLHAMNPRVQGEEL 104
```

```
Qy      51 FGLV-----ESSTPNAINPDDSSSEYRWYGYESYVRGHPYYKQFRVAHDLRVNLEGSRSYQ 105
Db       1 FGQVAYAADEKTVPNRIN--SNPEPPWYGTDSYSGRRLRV----HNLKVNLNGSKEYQ 52
```

```
Qy      51 FGLV-----ESSTPNAINPDDSSSEYRWYGYESYVRGHPYYKQFRVAHDLRVNLEGSRSYQ 105
              ||| |    |:|| ::||| :| ||:|  | | |:|||  | |||| :|||||||||| ||
Db       1 FGQVAYAADEKTVPNRIN--SNPEPPWYGTDSYSGRRLRV----HNLKVNLNGSKEYQ 52
```

```
Query Match             6.0%;  Score 236.5; DB 16;  Length 340;
Best Local Similarity  25.9%;  Pred. No. 2e-05;
Matches  96;  Conservative  54;  Mismatches  118;  Indels  103;  Gaps  16;

Qy     404 TTQNYAKFYYAKNKN--GSSQVYYCENADLKSPPD---SEDGGKTMTPDFTTGEVKYTHI 458
Db      25 TVVNGAKLTVTKNLDLVNSNALIPNTDFTFKIEPDTTVNEDGNKFKGVALNTPMTKVTY- 83

Qy     459 AGRDLFKYTVKPRDTDPDTFLKHIKKVIEKGYREKGQAIEYSGLTETQLRAATQLAIYF 518
Db      84 -------------TNSD---------------KGGSNTKTAEFDFSEVT------FEKPGVYYY 113

Qy     519 TDSAE-LDKDKLKDYHGFGDMNDSTLAVAKILVEYAQDSNPPQLTDLDFFIPNNNKYQSL 577
Db     114 KVTEEKIDKVPGVSY-------DTTSYTVQVHVLWNEEQKRPVAT------YI 153

Qy     578 IGTQWHPEDLVDIIRMEDKKEVIPV-------THNLTLRKTVTGLAGDRTKDFHFIELK 630
Db     154 VGYK--------EGSK--VPIQFKNSLDSTTLIVKKVSGTGGDRSKDFNFGLTLK 199

Qy     631 NNKQELLSQTVKTDKINLEFKDGKATI------NLKHGESLTLQGLPEGYSYLV 678
Db     200 ANQYYKASEKVMIEKTI--KGGQAPVQTEASIDQLYHFTLKDGESIKVTNLPVGVDYVV 256

Qy     679 KETD--SEGY--KVKVNSQEVA-----NATVSKTGITSDETLAFENNKEPVVPTGVDQK 728
Db     257 TEDDYKSEKYTNVEVSPQDGAVKNIAGNSTEQETSTDKDMTITFINKKDFEVPTGVAMT 316

Qy     729 INGYLALIVIA 739
Db     317 VAPYIALGIVA 327
```

```
Query Match             5.7%;  Score 224.5; DB 16;  Length 696;
Best Local Similarity  22.2%;  Pred. No. 0.00025;
Matches 166;  Conservative 102;  Mismatches 256;  Indels 225;  Gaps  42;
```

```
Qy    21  KNSKR--FTVTLVGVFLMIFALVTSMVGAKTVFGLVESSTPNAINPDSSSEYR------W  72
          |||  :  : |  :: |  :  :  :  ||
Db     4  KNSYKLSFLLSLTG---FILGLLLVFIGLSGV-SVGHAETRNGANKQGAFEIKKNKSQEE  59

Qy    73  YGYESYVRGHPYYKQFRVAHDLRVNLEGSR------SYQVYCPFNLKRAFPLGSDSSVKKW 126
          |  | :   ||   :  |        | :||| || | | ||
Db    60  YNYE-----VVDNRNILQDGEHKLEIKRVDGTGKTYQGFCFQLTKNFPTAQGVS-KKL  111

Qy   127  YKKHDGISTKFED----------YAMSPR-ITGDELNQKLRAVMYINGHPQNANGIMEGLEPL 177
          ||  |  |||  ::      ::| |   |    :| |
Db   112  YKK----LSSSDEETLKQYASKYTSNRRGDTSGNLKKQIAKVLTEGYPTNKSDWLNGLTEN  168

Qy   178  NAIRVTQEAVWYSDNAPISNPDDESFKRESESNLVSTSQLSLMRQALKQLIDPNLATRKMP 237
          |||| : | : |||||: |:|:   :  :|||: | |:  ::: |  : |
Db   169  EKIEVTQDDAIWYFTETTVPA--DRSYINRNVNS-----QKMKEVYQKLID---TIDID  216

Qy   238  KQVPDDFQLSIFESEDKGDKYNKGYQNLLSGGLVPTKPPTPGDPPMPPNQPQTTSVLIRK 297
          | ::  | :     | |    |     :||   :|  |:| | |:  |  |||
Db   217  KY--EDVQFDLFVPQD-------TNLQAVIS--VEPVIESLPWTSLKPIAQKDITAKKIWV  266

Qy   298  YAIGD----YSKLLEGATLQLTGDNVNSFQARVFSSNDIGERIELSDGTYTLTELNSPAG 353
             |||         |::|   :   | | :|| :  ||: | |||:| |  ||:| :||
Db   267  DAPKEKPIIYFKLYR----QLPGEK-------EVAVDDAELKQINSEGQ  304

Qy   354  YSIAEPITFKV---EAGKVYTIIDGKQIENPNKEIVEPYSVERAYNDF-EEFSVLTQNYA 409
             | |: || |     |::| :|  | ::||:|  |: || | ::|| ::||:| | |
Db   305  QEISVTWTNQLVTDEKGMAY-IYSVKEVDK-NGELLEP------KDVIKKEDGLTVIN--  354

Qy   410  KFYYAKNKNGSSQVVYCF-NADLKSPPDSEDGGKTMTPDFTTGEVKYTHIAGRD------ 462
             | | : | |  | |  :  | |: | |||  :| |  ||:  | | ::| |
Db   355  --TYVKPTSGHVDIEVTFGNGHIDITEDT------TPDIVSGENOMKQIEGEDSKPIDE  405

Qy   463  ------LFKYTV-------------KPRDT----------------DPDTF 478
                 | |:|                |  |                  | |
Db   406  VTENNLIEFGKNTMPGEEDGTNSNKYEEVEDSRPVDTLSGLSSEQGQSGDMTIEEDSATH  465

Qy   479  LKHIKKVIEKGYREKGQAIEY---SGLTETQLRAATQLAIYY---------FTDSAELDKDK 528
             | | : | : |  | ||  |     | |  |  | |  |          | |:|| | |
Db   466  IKFSKRDID-GKELAGATMELRDSSGKTISTWISDGQVKDFYLMPGKYTFVETAAPDGYE  524

Qy   529  LKDYHGFGDMND--------------STLAVAKILV--EYAQDSNPPQLTDLDFFIPNNNKYQSL  577
             | |  |  |:              :| | | |:    ||   | | :  | ||| | :: : |
Db   525  IATAITF-TVNEQGQVTVNGKAITKGDAHIVMVDAYKPTKGSGQVIDIEEKLPDE------  577

Qy   578  IGTQWHPEDLVDIIRMEDKK----------EVIPVTHNLTLRKTVTGLAGDR------  619
             |  |  |:   :  |                 ::| :::|  |
Db   578  --OGHSGSTTEI---EDSKSSDVIIGGQGQIVETTED-----TQTGMHGDSGCKTEVED  626

Qy   620  ---TKDFHFEIELKNNKQELLSQTVKTDK  645
             :  ||||  |:  |  | | ||  |  |  |
Db   627  TKLVQSFHFD---NKESESNSEIPKKDK  651
```